DOCUMENT RESUME

ED 415 285 TM 028 031

AUTHOR Bobrow, Warren
TITLE Development and Validation of a Personnel Assessment System for Selection of Tufting Machine Operators.
PUB DATE 1997-12-00
NOTE 14p.
PUB TYPE Reports -- Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Aptitude Tests; Decision Making; Employment Qualifications; Job Skills; *Machine Tool Operators; Mathematical Aptitude; Mechanical Skills; *Occupational Tests; Personality Assessment; *Personality Traits; *Personnel Selection; Predictive Measurement; Sex Differences; Sex Discrimination; Test Construction; Test Use; Test Validity; Thinking Skills; Visual Acuity; *Vocational Evaluation

ABSTRACT

This paper outlines the results of a test validation project for entry-level positions in the tufting areas in a carpet factory. Carpet factory administrators thought that new methods of hiring associates would help improve overall system productivity. A job analysis was conducted through review of job materials and focus groups with employees and supervisors. Critical knowledge, skills, and abilities were identified for the job at present and in the future. The following dimensions were found to be critical to job performance: (1) decision making ability; (2) arithmetic ability; (3) conscientiousness; (4) visual acuity; (5) mechanical aptitude; and (6) agreeableness. Instruments, which included tests of abstract reasoning, the ability to spot flaws in a series of pictures, a mathematics test, a visual speed and acuity test, a test of mechanical aptitude (with little or no gender bias), and an attitude questionnaire were used in pilot testing with 96 current employees, whose productivity was assessed to provide a measure of criterion validity. Pilot testing indicated that the assembled test battery was valid for the selection of tufting machine operators without adverse impact against women. The test construction and validation procedures outlined in this project are comparable and could easily be transferred to those used in educational assessment. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 415 285

TM

Development and Validation of a Personnel Assessment System for Selection of Tufting Machine Operators

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Warren Bobrow
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Prepared by:
Warren Bobrow, Ph.D.
The Context Group
239 Winchester Street
Newton, Massachusetts 02161
(617) 630-1021

December, 1997

TM028031

BEST COPY AVAILABLE

Development and Validation of a Personnel Assessment System for Selection of Tufting Machine Operators

Warren Bobrow, Ph.D.
The Context Group

This paper outlines the results of a test validation project for entry level positions in the tufting areas in a carpet factory ("Company"). The Company was seeking to improve the overall performance of the business and the productivity of their tufting machine operators (tufting machines make carpet). It was felt that, along with other changes in the organization, developing new methods of hiring associates would help improve overall productivity. The decision was made to develop and implement a valid selection system to fill job openings.

Job Analysis

A job analysis was conducted by first reviewing existing job materials (job descriptions, training manuals, etc.). This review was augmented by focus groups that were conducted with incumbents from each classification and their direct supervision. This process was used to determine what the critical knowledge, skills, abilities and personal characteristics (KSAPs) to do the jobs now are as well as identify those that would be needed in the future.

Based upon the job analysis survey and focus groups, the following dimensions were found to be critical to the performance of the jobs:

1. Decision Making Ability

Being able to solve a variety of problems effectively. Taking decisive action and communicating with confidence. Recognizing when additional information is needed and looking for that information. Being able to make high quality decisions based on analysis of the available information.

2. Arithmetic Ability

Ability to solve mathematical problems as they relate to counting and planning.

3. *Conscientiousness*

The desire to be concerned about the effects of one's actions on other people and the organization. Being trustworthy enough to attend work regularly and on time.

4. *Visual Acuity*

The ability to visually check for errors. This acuity includes seeing errors in written material (e.g., material codes and lot numbers) and carpet materials and yarn.

5. *Mechanical Aptitude*

Understanding of basic mechanical principles and how they apply to industrial and ordinary settings.

6. *Agreeableness*

The desire to get along with others, rather than challenge them. Also includes trusting others' judgment and being willing to work with others.

Implications of the Job Analysis Results for Test Development

The results of the job analysis suggest that there were several important KSAPs that needed to be measured during the assessment process. All selection instruments were selected to tap into those KSAPs. However, it was important for the company that the final selection system not have adverse impact against women (the company exists in a racially homogeneous area where there are very few minorities). Previous research and experience indicated that other published measures of mechanical ability (e.g., The Bennett Mechanical Test) do have adverse impact against women. For example, the gender difference between 12th grade boys and girls taking the DAT Mechanical Reasoning Test, Forms L and M is 1.22 standard deviation units in favor of boys (Bennett, 1994, Table 4.7, page 30). Therefore, another test of mechanical ability, the Wiesen Test of Mechanical Aptitude (WTMA), was used in the study. The questions in the WTMA are based on everyday objects and situations, and so the WTMA was expected to have less adverse impact on women than traditional tests of mechanical aptitude.

Development of Selection Instruments

The following tests were used in the validation study:

- *Abstract Reasoning Test* (Resource Associates). This 40 item test measures analytical and problem solving abilities using non-language based items.
- *FACT-1A Inspection* (SRA). This 150 item test measures a person's ability to quickly and accurate identify flaws in a series of object pictures.
- *Math Test* (The Context Group). This is a 25 item test that measures basic math knowledge and application of mathematical concepts.

- *Visual Speed and Accuracy Test* (The Context Group). This is a 135 item test that measures a person's ability to correctly match pairs of number/letter combinations.

- *Wiesen Test of Mechanical Aptitude* (Applied Personnel Research; Newton, Massachusetts). This test contains 60 items that measure a person's ability to understand and apply basic mechanical principles. The WTMA yields an overall score and also provides information on 11 subscales that are designed for research. Three of the subscales reflect the type of object used in the question, and 8 of the subscales reflect the physical/mechanical principle. The subscales are listed below.

  General environmental areas or types of objects:

  SubO1: Kitchen objects (e.g., can opener, steak knife, ice)
  SubO2: Non-kitchen household objects (e.g., flashlight, chair, step-stool, mirror)
  SubO3: Other everyday objects (e.g., ladder, car, bus, bridge, shadows)

  Broad classes of physical/mechanical principles:

  SubP1: The six basic machines (lever, wheel and axle, pulley, inclined plane, wedge, screw)
  SubP2: Movement of simple and complex objects (e.g., fluid flow, inertia, wheels and gears)
  SubP3: Center of gravity, and gravity
  SubP4: Basic electricity/electronics (e.g., flashlight batteries)
  SubP5: Transfer of heat (e.g., radiant heat, conduction of heat, as in cooking)
  SubP6: Basic physical properties of matter and materials (e.g., weight, strength)
  SubP7: Miscellaneous (e.g., optics, friction, hinges and locks)
  SubP8. Academic (most likely learned in school)

- *Attitude Questionnaire* (The Context Group). This 152 item personality test was composed of items measuring the constructs of *Agreeableness* and *Conscientiousness*.

  Facets of Agreeableness included:

  Trust
  Altruism
  Compliance
  Openness to working with others

Facets of Conscientiousness included:

Competence
Order
Dutifulness
Achievement Striving
Self-Discipline
Deliberation

Results of Pilot Testing

Pilot testing was optional for associates and was conducted for all the written tests. Associates who have been on their jobs for 90 days or more were asked to participate. Pilot testing was conducted in groups of between 10 and 25 during regular shift hours. The pilot testing was conducted by the Company's human resources representatives after they received training on proper methods and procedures for giving the tests. Tests were not scored by the Company to ensure confidentiality. Participants also had the opportunity to request test feedback.

Descriptive Statistics of the Written Test Scores

Descriptive statistics of the test scores are summarized in Table 1. Intercorrelations and reliability of test scores are shown in Table 2.

Criterion Development

Productivity was measured through a performance appraisal form designed exclusively for this research project. They were told these ratings were for the validation study only and the scores would not affect the associates in any way. Supervisors and managers received training on how to avoid common rater errors (halo, central tendency and leniency) and keep the ratings confidential. Analysis of the ratings indicated that they were internally consistent (α=.96) and reliable between raters (α=.80). As such, the 20 items were summed to form a final rating. This information was gathered for all entry level associates who had been on the job at least 90 days. Due to significant differences in the average ratings given by supervisors, performance appraisal scores were standardized (mean of 0, standard deviation of 1) for all subsequent analyses.

Table 1
Descriptive Statistics for the Test Scores

Variable	N	Minimum	Maximum	Mean	Standard Deviation
Abstract Reasoning	94	-5.00	26.75	8.55	8.34
Math	96	.00	25.00	11.20	5.30
Visual Speed and Accuracy	95	42.00	135.00	103.77	22.43
Inspection	96	8.00	87.00	55.81	16.24
WTMA	96	3.30	86.70	64.05	13.73
SUBO1	96	5.00	90.00	67.97	14.10
SUBO2	96	5.00	95.00	65.42	16.38
SUBO3	96	.00	95.00	58.75	17.52
SUBP1	96	.00	100.00	67.32	22.82
SUBP2	96	.00	100.00	66.41	15.10
SUBP3	96	.00	100.00	72.01	20.82
SUBP4	96	.00	100.00	64.57	22.88
SUBP5	96	.00	100.00	71.56	17.88
SUBP6	96	.00	100.00	64.19	19.50
SUBP7	96	.00	100.00	63.93	23.36
SUBP8	96	.00	100.00	41.53	18.06
Order	96	1.00	5.00	3.46	.69
Openness to Experience	96	2.00	3.93	3.16	.40
Competence	96	1.00	5.00	3.47	.66
Achievement Striving	96	1.00	4.86	2.82	.75
Trust	96	1.63	5.00	3.61	.65
Compliance	96	1.00	5.00	3.65	.68
Altruism	96	1.00	5.00	3.89	.68
Dutifulness	96	2.00	5.00	3.75	.60
Deliberation	96	1.63	4.75	3.25	.68
Self-Discipline	96	2.00	5.00	3.68	.66

Table 2
Inter-Correlations Between Test Scores

Reliabilities (coefficient α) are in bold in the diagonal

	Abstract Reasoning	Math	Visual Speed and Accuracy	Inspection	WTMA	SubO1	SubO2	SubO3	SubP1	SubP2	SubP3	SubP4	SubP5	SubP6	SubP7	SubP8	Order	Openness to Experience	Competence	Achievement Striving	Trust	Compliance	Altruism	Dutifulness	Deliberation	Self-Discipline
Abstract Reasoning	**.93**																									
Math	.50	**.70**																								
Visual Speed and Accuracy	.52	.70	**.70**																							
Inspection	.48	.44	.56	**.77**																						
WTMA	.45	.36	.49	.43	**.84**																					
SubO1	.23	.23	.42	.29	.83	**.53**																				
SubO2	.39	.38	.43	.38	.87	.61	**.67**																			
SubO3	.50	.31	.42	.42	.86	.57	.61	**.69**																		
SubP1	.39	.26	.35	.40	.79	.63	.67	.71	**.63**																	
SubP2	.26	.29	.36	.36	.70	.51	.61	.65	.40	**.12**																
SubP3	.20	.30	.48	.18	.66	.65	.49	.57	.53	.43	**.51**															
SubP4	.34	.33	.38	.35	.60	.45	.58	.51	.39	.44	.20	**.49**														
SubP5	.32	.21	.24	.19	.67	.60	.62	.52	.42	.42	.47	.28	**.28**													
SubP6	.19	.11	.25	.28	.69	.53	.62	.62	.44	.47	.28	.43	.38	**.21**												
SubP7	.42	.31	.37	.28	.74	.64	.66	.59	.54	.38	.45	.32	.46	.44	**.63**											
SubP8	.25	.12	.24	.28	.53	.46	.41	.51	.43	.38	.23	.17	.31	.30	.23	**.04**										
Order	.13	-.01	.07	-.00	.37	.34	.34	.29	.19	.19	.22	.20	.30	.37	.30	.25	**.60**									
Openness to Experience	-.12	-.13	-.22	-.06	-.14	-.10	-.09	-.17	-.16	-.08	-.16	-.09	-.09	-.09	-.02	-.06	-.00	**.45**								
Competence	.10	.07	.13	.11	.31	.33	.26	.22	.16	.15	.24	.14	.23	.33	.28	.14	.57	-.02	**.65**							
Achievement Striving	.07	-.04	.01	.08	.04	.00	.03	.07	-.01	.03	-.02	.05	.06	.05	.02	.07	.38	.17	.50	**.59**						
Trust	.00	.07	-.03	.00	.10	.11	.21	-.04	.03	.08	.12	-.00	.12	.05	.22	-.10	.13	.24	.12	.01	**.70**					
Compliance	-.17	-.08	.03	-.07	.12	.19	.09	.04	.12	.01	.26	-.21	.08	.21	.16	-.06	.26	-.23	.18	-.17	.10	**.71**				
Altruism	.10	.20	.28	.17	.35	.33	.33	.25	.20	.39	.24	.15	.19	.42	.25	.07	.29	-.08	.37	.00	.28	.38	**.67**			
Dutifulness	-.06	-.03	.02	-.06	.12	.14	.13	.05	.03	.11	.14	-.04	.07	.26	.13	-.06	.44	.00	.32	.28	.18	.37	.19	**.48**		
Deliberation	-.08	-.03	-.11	-.02	.09	.14	.06	.04	.02	.05	.00	.06	.04	.18	-.00	.13	.57	.05	.49	.29	.07	.05	.06	.36	**.69**	
Self-Discipline	.11	.02	.14	.16	.30	.31	.27	.21	.06	.16	.20	.21	.23	.40	.28	.07	.60	-.09	.66	.41	.14	.19	.33	.56	.51	**.67**

Note: Given that incumbents are not a random sample, the true reliabilities of the tests may be higher than what is represented in the table.

BEST COPY AVAILABLE

Criterion-Related Validation Analysis

Table 3 shows the correlation between test scores and the performance appraisal ratings for the ability measures, with experience partialled out. (The zero-order *r*'s were within about .05 of these partial correlations.) The results indicate that scores on the FACT and WTMA were related to job performance. Due to its higher reliability, the overall WTMA score was retained. Due to the high intercorrelations between the ability tests, regression analysis indicated that adding any valid test after another valid test in the equation did not significantly increase the R^2. However, adding the Achievement Striving test after the WTMA did significantly increase the R^2 (see Tables 4 and 5).

Adverse Impact of the Final Test Battery

For this analysis, the predictor scores of different protected groups were compared. This is especially important of the WTMA because mechanical ability tests have traditionally had adverse impact against women (Wiesen, 1997). The group comparisons were based on mean scores and distribution scores of groups. The *Total* score represents a unit weighted composite of the standardized scores of the two predictors. This would be the score that would be used for making selection decisions. We determined in advance that a predictor would not be likely to have adverse impact if group means were within ½ standard deviation of each other. The adverse impact analysis for gender is presented below in Table 6.

Table 3
Partial Correlations Between Ability Tests and Criterion Measure (n=95)

	AR	Math	VSA	FACT	WTMA	SUBO1	SUBO2	SUBO3	SUBP1	SUBP2	SUBP3	SUBP4	SUBP5	SUBP6	SUBP7	SUBP8
Performance Ratings	-.07	.04	.07	.19*	.20*	.21*	.22*	.10	.14	.22*	.03	.15	.05	.14	.19*	.18*

Note: * $p < .05$ (one-tailed)
Correlations are not adjusted for range restriction or unreliability.
AR = Abstract Reasoning, VSA = Visual Speed and Accuracy

Table 4
Partial Correlations Between Personality Tests and Criterion Measure (n=95)

	Order	Openness	Competence	Achievement Striving	Trust	Compliance	Altruism	Dutifulness	Deliberateness	Self-Discipline
Performance Ratings	.06	-.02	.11	.20*	.10	.02	.03	.09	-.06	.16

Note: * $p < .05$ (one-tailed)
Correlations are not adjusted for range restriction.

BEST COPY AVAILABLE

Table 5
Regression Analysis of Valid Predictors

	R	R^2	R^2 Change	F Change
WTMA	.20	.04	.04	3.90* (1,93)
WTMA + Achievement Striving	.28	.09	.05	3.67* (1,92)

Note: * $p < .05$

Table 6
Gender Adverse Impact Analysis of the Predictors

	Gender	*n*	Mean	S.D.	Partial *r*
WTMA	Female	70	64.36	11.82	.22*
	Male	26	63.24	18.15	.29
	Total	96	64.05	13.73	.21*
Achievement Striving	Female	70	2.73	.76	.17
	Male	26	3.03	.69	.19
		96	2.82	.75	.21*
Total	Female	70	-.09	.73	.25*
	Male	26	.07	.79	.35*
		96	-.04	.71	.29**

Note: * $p < .05$, ** $p < .01$ (one-tailed)

As can be seen in Table 6, the mean scores for men and women are within ¼ standard deviation on the composite, and within ½ standard deviation for both of its two components, so the predetermined standard we set in this regard was met. This indicates that it is unlikely that the tests will have adverse impact against women when used on applicant populations. Because the validation population was nearly all white, it is impossible to examine the adverse impact of these tests against racial minorities in this sample.

It may be that the WTMA did not have had adverse impact because the principles measured in the test reflect everyday occurrences as shown in a sample item below rather than concepts that one would only learn through working on machinery.

Conclusion

A valid test battery consisting of the WTMA and a test of achievement striving was designed for tufting machine operators. This test battery demonstrated criterion-related validity against supervisory ratings. Just as important, the WTMA was shown not to have adverse impact against women. This is unusual for a test of mechanical aptitude.

REFERENCES

Bennett, G.K. (1994) *Manual BMCT: Bennett Mechanical Comprehension Test (2nd ed.)* San Antonio: The Psychological Corporation.

Wiesen, J.P. (1997) *Technical Manual for the Wiesen Test of Mechanical Aptitude* (WTMA). Newton, Massachusetts: Applied Personnel Research.